Naval Submarine Medical Research Laboratory

NSMRL REPORT 1181 2 SEPTEMBER 92

TRIAL OF A COMPUTER BASED PROGRAM FOR THE DIAGNOSIS OF ABDOMINAL PAIN IN MALES

by

P. L. Perrotta
and
D. M. Stetson

Released by
R. G. Walter, CAPT, DC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

TRIAL OF A COMPUTER BASED PROGRAM FOR THE DIAGNOSIS OF ABDOMINAL PAIN IN MALES

by

Peter L. Perrotta and Douglas M. Stetson

Naval Submarine Medical Research Laboratory
Report 1181

Naval Medical Research and Development Command
Work Unit 63706N M0095.005-5010

Released by

R. G. Walter
Commanding Officer

SUMMARY PAGE

PROBLEM:

To verify the knowledge base of the NSMRL abdominal pain program in previously healthy males.

FINDINGS:

Overall diagnostic accuracy of the program was found to be 69% compared to the 80% accuracy rate of emergency room physicians. Sensitivity and specificity for distinguishing surgical from non-surgical cases was 56% and 85% respectively. Additional measures of program performance are presented.

APPLICATION:

These results can assist in the decision whether to implement this diagnostic program for fleet-wide use. Recommendations are given for additional efforts in medical diagnostic software.

ADMINISTRATIVE INFORMATION

This investigation was conducted under Naval Medical Research Development Command Research Work Unit 63706N-M0095.005-5010. The views expressed in this report are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the U. S. Government. This report was approved for publication on 2 September 1992 and designated Naval Submarine Medical Research Report 1181.

Abstract

This report presents and evaluates data collected in 1988 in an effort to verify the NSMRL abdominal pain diagnostic program. Overall diagnostic accuracy of the program was found to be 69% compared to the 80% accuracy rate of emergency room physicians. Sensitivity and specificity for distinguishing surgical from non-surgical cases was 56% and 85% respectively. Additional measures of performance of the abdominal pain program are presented along with the limitations of the data set and recommendations for future validation efforts.

[Blank Page]

TRIAL OF A COMPUTER BASED PROGRAM FOR THE DIAGNOSIS OF ABDOMINAL PAIN IN MALES

Background

The Naval Submarine Medical Research Laboratory (NSMRL) developed several computer based diagnostic programs between 1970 and 1988 to be used as diagnostic aids for corpsman and health care practitioners in remote duty stations. One of these (1,2), based on a Bayesian knowledge base developed in England (3), assisted in the diagnosis of acute abdominal pain. The original algorithm was then modified for an active duty population (young healthy males presenting within 48 hours of illness) (4). A computer program incorporating this modified knowledge base was developed and tested by the knowledge-base authors in England using locally obtained clinical data as well as data obtained from a Navy hospital in 1980.

In an effort to conduct independent verification of the program's performance in the hands of submarine independent duty hospital corpsmen, NSMRL undertook a study collecting abdominal pain case data from submarines at sea. Five years of data collection at sea yielded very few cases, largely because submarine sailors are generally in good health and carefully screened prior to deployment (5).

To collect a larger number of cases, the laboratory undertook a prospective study in 1988 collecting data from the emergency rooms of two Naval hospitals. Although the Navy hospital emergency rooms see large numbers of patients, the minority are Naval personnel on active duty. In order to gather a reasonable number of cases in a shorter period of time, clinical data was collected from family members presenting to the emergency room as well as from active duty personnel. From this group of patients, only information about males in the same age range as submariners and whose medical history would not disqualify them from submarine duty would be used as test data for evaluation of the abdominal pain program.

Methods

Data Collection

With the approval of the hospital commanders and the individual patients involved in the study, specially hired clerks collected information from patients during emergency room visits. All were college students with a background in biological sciences. The clerks avoided delaying care of patients with acute medical problems and burdening the hospital staff. They selected for interview all patients who listed abdominal pain as part of their reason for seeking care when they presented themselves to the emergency room staff. The clerks covered the time period from 8:00 AM to 12:00 midnight.

NSMRL investigators trained the

clerks in use of the abdominal pain program on portable computers and provided the clerks with blank data forms for use in collecting data from individual patients. When a patient complaining of abdominal pain visited the emergency room, a clerk would collect medical history information from the patient directly, recording the results on the data forms. The clerk would then accompany the patient and observe the interview and examination by the emergency room physician. If at any time a patient requested the clerks not be present for the interview or examination, the clerks would leave and not include this information in the study database. By observing the interview and examination conducted by the physician, the clerk completed the data collection form. On those occasions where urgency prevented the clerk from conducting independent history taking, the history was gathered from the observation process. When the information gathered by observation was insufficient to complete the data form, the clerks would query the examining physician for the missing information. If the physician was too busy to provide information, the clerk would review the written emergency treatment record to fill in missing items. Cases which had missing items were discarded.

At a later time the clerk entered the information from the data forms into the abdominal pain computer program to create a database showing patient information and computer results. Those cases which had insufficient information to make a diagnosis were not included in the computer database. In addition, the clerks kept a written log showing the patients and the diagnoses assigned at the emergency room along with a contact telephone number.

Several weeks to months after the patients were seen in the emergency rooms, a clerk or an investigator contacted each one by telephone to inquire whether subsequent events had cast doubt on the emergency room diagnosis. If the patient was not seen again for the same problem, the emergency room diagnosis was taken to be confirmed. If a patient was admitted to the hospital the discharge diagnosis was recorded as the confirmed diagnosis for the emergency room visit. The written log was annotated to reflect the confirmed diagnoses and patients lost to follow-up.

Data Review

Case acceptance criteria

Prior to evaluation of any data, acceptance and exclusion criteria were determined. The criteria were selected to mirror those used by the developer of the knowledge base upon which the abdominal pain diagnostic system rests (4). These criteria were: male patients age 17-50, who presented to the emergency room with a complaint of abdominal pain and who had no chronic illness which would have been disqualifying for submarine service.

Case categorization criteria

The abdominal pain diagnostic system considers only six diagnoses (Table 1: appendicitis, perforated duodenal ulcer, small bowel obstruction, cholecystitis, renal colic, and non-specific abdominal pain). The

program classifies all cases submitted into one of these categories. de Dombal (4) previously described how the program categorizes less common conditions based on his experience. Categorization of common conditions rests on professional judgment of the reviewers. For example, viral gastroenteritis is placed in the non-specific category.

Table 1. Diagnoses Considered by Abdominal Pain Program

Diagnosis	Abbreviation
Non-Specific Abdominal Pain	NSAP
Appendicitis	APPY
Cholecystitis	CHOL
Perforated Duodenal Ulcer	PDU
Renal Colic	RENC
Small Bowel Obstruction	SBO

Case grouping

In this analysis, diagnoses were grouped according to usual treatment requirements. For the present evaluation, diagnoses of appendicitis, perforated duodenal ulcer, and small bowel obstruction were categorized as surgical; diagnoses of renal colic, cholecystitis, and non-specific abdominal pain were categorized as non-surgical. This distinction was made because the most important decision to be made aboard submarines is often the decision to seek medical evacuation and definitive care.

Data verification

Data analysis was conducted based on a composite database comprising elements of other files. To ensure its accuracy, and before relying on derivative databases, the source documents (log books prepared at Portsmouth and San Diego) were reviewed to ensure accurate transcription. Each male case in the 17-50 year old age range was reviewed in a written log. The emergency room and confirmed diagnoses were verified when present, and their absence was specifically noted in the composite database. Each log diagnosis was categorized into one of the six diagnostic categories considered by the program. Rare diagnoses were categorized according to de Dombal's guidance as previously described. Cases regarded as chronic or occurring in patients having medical conditions incompatible with submarine service were excluded.

Database Management

Initially, a computer database was prepared on site at each study hospital. This record contained the elements of the patient descriptions recorded by the clerks. These

databases were examined to ensure that derivative records, used to create the composite database, were complete and accurate. When questions about the source data arose, these source files could often be used to ensure accuracy of the results. Data taken directly from the site databases were used as entries in the abdominal pain diagnostic program to observe whether the program's output was faithfully recorded in derivative databases.

Statistical Analysis

As noted, only cases identified as male, age 17-50 years, presenting to the emergency room with abdominal pain and not evincing chronic conditions which would have disqualified them from submarine duty were analyzed.

All statistical work was done using SPSS-PC, drawing on the cases contained in the composite database. The test data set was subjected to frequency analysis of individual diagnoses, distribution of computer generated diagnostic frequencies against individual diagnoses, and cross tabulation of computer diagnoses against either the final diagnosis for each case or the emergency room diagnosis. For purposes of comparison, the emergency room diagnoses were also compared with the final diagnoses when both were known.

Based on the cross tabulation data, diagnostic accuracy (percentage of diagnoses "correct") and sensitivity/specificity for each diagnosis was calculated. Chi-square analysis was performed including calculation of Cramer's V coefficient, Goodman and Kruskal's tau (percentage reduction in error), and Cohen's kappa (measure of agreement).

Results

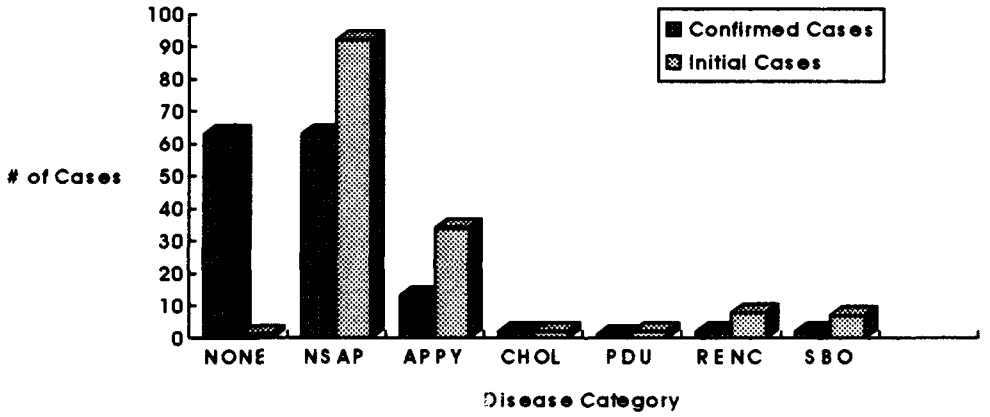
The initial data collected at the two Naval hospitals yielded 616 total cases (Table 2). Of these, less than one-half were male (34%). Additional cases were excluded in accordance with the pre-established criteria (age 17-50, presenting with abdominal pain, and with no illness disqualifying for submarine duty). A total of 146 cases remained for analysis. All remaining results refer to the test set of males, age 17-50, with no chronic, submarine disqualifying illness.

Table 2. Preliminary Categorization of Test Data

	Number of Cases	Percent of Cases
Total Collected (both sites)	616	100
Male	208	34
Age 17-50	171	28
Abdominal Presentation	153	25
No disqualifying illness	146	24

Figure 1 presents the disease distribution of the cases analyzed. Confirmed diagnoses are reported for those cases with adequate follow-up. Initial diagnosis refers to the diagnosis made by the emergency room physician during the initial patient encounter. The largest category of cases (43.2%) was Non-Specific Abdominal Pain, the remaining disease categories included few cases. The category "NONE" represents cases that had no emergency room diagnosis or no confirmed diagnosis.

Figure 1
Distribution of Cases



Diagnostic Accuracy

The overall accuracy of the abdominal pain diagnostic system was determined by dividing the number of cases the computer correctly categorized by the number of cases attempted. Both the confirmed diagnosis and the emergency room diagnosis were separately considered as the correct diagnosis. The accuracy of the diagnostic program was 69% when compared to either the final diagnosis or the emergency room diagnosis (Table 3). By comparison, the accuracy of the emergency room physician on initial examination was 80%.

Table 3. Diagnostic Accuracy of Program

Comparison Bases	Accuracy (%)
Computer Dx -vs- Final Dx	69
Computer Dx -vs- Emergency Room Dx	69
ER Physician Dx -vs- Final Dx	80

Sensitivity and Specificity

The sensitivity of the abdominal pain diagnostic system for the diagnosis of appendicitis was determined because appendicitis is the most common abdominal medical condition requiring evacuation and surgical intervention. Sensitivity is defined as the true positive rate for the diagnosis (the fraction of those patients with a diagnosis of appendicitis correctly identified by the computer as having the diagnosis). The specificity of the abdominal pain diagnostic system is defined as the true negative rate for a diagnosis (the fraction of those patients not diagnosed with a particular illness who were correctly identified as not having that illness). For the diagnosis of appendicitis the computer program's sensitivity and specificity were 46% and 86% respectively (Table 4). Emergency room physicians sensitivity was 83% for comparison.

Table 4. Sensitivity and Specificity for Diagnosis of Appendicitis

	Sensitivity	Specificity
Computer vs. Confirmed Dx	0.46	0.86
Computer vs. Emergency Room Dx	0.50	0.94
Physician vs. Confirmed Dx	0.83	0.83

Cross tabulation data

Cross tabulations were prepared comparing the computer diagnostic performance with confirmed diagnoses and emergency room diagnoses (Tables 5,6,7). Several statistics describe the relationship between the observed results (from the computer) and the expected results (either the confirmed diagnoses or the emergency room diagnoses) were calculated. Chi-square with Cramer's V coefficient was used to account for sample size when comparing group differences. As Cramer's V coefficient approaches 1 the probability that chance accounts for the differences between samples

falls. Calculation of proportional reduction of error (the Goodman and Kruskal tau) and a measure of agreement (Cohen's kappa) were calculated (Table 8).

Tau is used as a measure of the benefit of using the diagnostic program over methods based on knowledge of the underlying disease prevalence and varies between 0 and 1. If there is no benefit from using the information about distribution of test categories over outcome categories, tau is zero. The significance reported with tau is the probability that tau is zero.

Cohen's kappa is an additional method of evaluating the level of agreement between the computer program and the emergency room physicians. The program and emergency room physicians categorized 80% (66 out of 82) of the cases identically but this calculation does not account that there will be some fraction of agreement even if both groups assigned diagnoses at random. The random agreement proportion is eliminated by noting the fraction of cases each observer places in each category and determining how many cases would be randomly assigned to the same category on that basis. Kappa ranges from 0 to 1, with 1 being associated with total agreement between the program and the physicians. Comparing the abdominal pain program with emergency room physicians, kappa is 56%.

Table 5. Computer Diagnosis vs. Confirmed Diagnosis

Final Dx / Computer Dx	NSAP	APPY	CHOL	PDU	RENC	SBO	Totals
NSAP	50	6	1	1	1	2	61
APPY	10	6	0	0	0	0	16
CHOL	0	0	0	0	0	0	0
PDU	0	0	0	0	0	0	0
RENC	3	0	1	0	1	0	5
SBO	0	1	0	0	0	0	1
Totals	63	13	2	1	2	2	83

Table 6. Computer Diagnosis vs. Emergency Room Diagnosis

Initial Dx	NSAP	APPY	CHOL	PDU	RENC	SBO	Totals
Computer Dx							
NSAP	79	16	2	1	4	7	109
APPY	7	17	0	0	0	0	24
CHOL	0	0	0	0	0	0	0
PDU	0	0	0	0	0	0	0
RENC	4	0	0	1	4	0	9
SBO	1	1	0	0	0	0	2
Totals	91	34	2	2	8	7	144

Table 7. Physicians Diagnosis vs. Confirmed Diagnosis

Final Dx	NSAP	APPY	CHOL	PDU	RENC	SBO	Totals
Physician's Dx							
NSAP	52	1	2	0	0	0	55
APPY	9	10	0	0	1	0	20
CHOL	0	0	0	0	0	0	0
PDU	0	0	0	1	0	0	1
RENC	2	1	0	0	1	0	4
SBO	0	0	0	0	0	2	2
Totals	63	12	2	1	2	2	82

Table 8. Statistical Summary

	Chi-square	Significance	Cramer's V	tau	kappa
Computer vs. Final Dx	28	0.02	0.33	0.10	-
Computer vs. ER Dx	71	-	0.28	0.08	0.36
Physician vs. Final Dx	203	-	0.78	0.39	0.56

Grouped Data (Surgical versus Non-surgical Cases)

Similar calculations were performed after grouping cases based on therapeutic implication into either surgical (appendicitis, perforated ulcer, small bowel obstruction) or non-surgical (non-specific abdominal pain, renal colic, cholecystitis) (Table 9). Diagnostic accuracy of the program

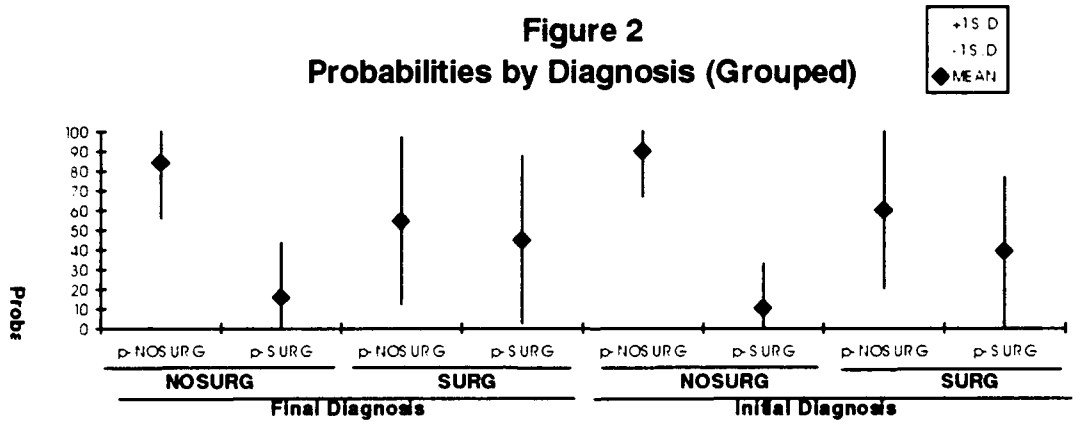
improved to 77% when data was grouped but sensitivity of the computer program for surgical diagnoses remained low at 43% (reflecting that most of the cases correctly categorized as "surgical" were appendicitis cases).

Table 9. Sensitivity and Specificity for Diagnosis of Surgical Cases

	Accuracy	Sensitivity	Specificity
Computer vs. Confirmed Dx	77%	0.43	0.85
Computer vs. Emerg. Room Dx	77%	0.41	0.92
ER Physician vs. Confirmed Dx	85%	0.86	0.85

Computer generated probabilities are compared in Figure 2 for the grouped data. Both the final confirmed diagnosis and emergency room diagnosis are considered. The mean computer-generated probability is indicated as well as +/- one standard deviation. For non-surgical cases there is little overlap in assigned probability for surgery required or no surgery required. But for surgical cases the overlap is substantial, indicating that the program generally categorizes non-surgical cases properly, but often mischaracterizes surgical cases as non-surgical.

Figure 2
Probabilities by Diagnosis (Grouped)



Discussion

Comparison to de Dombal's Findings

The de Dombal final report (4) described testing of the delivered Bayesian matrix. Three types of data were used to test the matrix: (a) cases which had been used to create the knowledge base matrix (b) additional cases from de Dombal's data file in England and (c) cases collected from Naval Hospital San Diego (cases different than those of this study). Testing was conducted against the original knowledge base

(3) and a knowledge base modified for the submarine community (4). These modifications included assigning equal conditional probabilities to all vital sign data and to the findings of rigidity and bowel sounds, and developing new conditional probabilities for signs and symptoms in a male patient population with no chronic conditions who had presented for care for abdominal pain beginning within the past 12 hours. The knowledge base was most effective when prior probabilities were also specified. Appendix A lists the prior probabilities used by the program and the conditional probabilities for each piece of clinical information. The overall accuracy of the abdominal pain diagnostic system reported in de Dombal's final report when tested against Navy data was 73%. The Navy case data used (Table 10) consisted of 141 cases and included few instances of diagnoses other than NSAP and APPY . de Dombal calculated accuracy as number of cases "correct" out of total tested. In the group above, the diagnoses of NSAP and DYSPepsia were considered equivalent and grouped together. The test set used in our analysis is similar in size (145) and contains a majority of NSAP diagnoses, a smaller number of appendicitis cases, and few cases of the remaining diagnoses. These results are comparable to those obtained with the present data set, where overall accuracy is 69%.

Table 10. Summary of de Dombal's Final Report

Final Dx	APPY	NSAP	DYSP	RENC	CHOL	SBO	DVRT	OTHER
Computer Dx								
APPY	25	18	1	0	0	1	0	2
NSAP	2	49	3	0	2	0	0	1
DYSP [a]	0	8	10	0	2	0	1	0
RENC	0	1	0	4	0	0	0	0
CHOL	0	1	0	0	2	0	0	1
PDU	0	1	0	0	0	0	0	0
SBO	0	2	1	0	0	3	0	0
Totals	27	80	15	4	6	4	1	4

a. The NSMRL program sums the probability of dyspepsia and NSAP and presents the total as NSAP in its output.

There are differences between the sensitivity and specificity for the diagnosis of appendicitis between de Dombal's report of the older San Diego data and the present test set (Table 11). The program's sensitivity in diagnosing cases of appendicitis was substantially better in the prior test set (92%). We know of nothing to explain this difference.

Table 11. Comparison With Previously Obtained Naval Data For Appendicitis

	Sensitivity	Specificity
Previous Data	0.92	0.80
Report Data	0.46	0.86

Limitations of the data

There are two major differences between the data in de Dombal's analysis and the present test set. Firstly, the source data was collected differently. de Dombal reports that hospital corpsmen collected the data he used. This suggests the data was collected by the person taking the history and performing the examination. The data of the present test set was collected by a third party with limited medical experience observing physicians. Secondly, the diagnostic probabilities were created by programs using different prior probabilities. The differing prior probabilities may interfere with the appendicitis diagnosis.

Another problem, extending to both data sets, is the inadequate numbers of cases in any but the NSAP category. Because of difficulties in verification of final diagnoses a large number of the cases collected in 1988 were excluded. A small number of additional cases had to be excluded because of uncertainty about the actual data included in the original files.

Recommendations for Further Study

Based on these results, it is not possible to access the clinical adequacy of the NSMRL abdominal pain diagnostic program. Certainly the program performed poorly with the cases of appendicitis (sensitivity of 46%). The program also had difficulty distinguishing surgical from non-surgical cases. Without the presence of pre-established criteria for acceptance or rejection of a diagnostic system it is difficult to make a definitive statement regarding the suitability of a program for clinical use. Currently, there are no accepted standards of performance for medical diagnostic systems prior to deployment in the Navy.

Knowledge verification may take many forms. These include review by experienced practitioners and trial against accumulated prospective or retrospective cases. A "gold-standard" test of verification against which other methods could be compared would require the testing of the system against cases prospectively gathered as was attempted in this study. Certain goals should be identified prior to data collection. This would include the minimum number of cases that should be collected for each diagnosis considered by the program. It is difficult to estimate the number necessary, but it is likely that a minimum of 20 to 30 cases of the less common illnesses would be desired. Once collected, these prospective cases could also be used to alter the existing knowledge base by changing

either the conditional probabilities or the disease prevalence information (prior probabilities of disease) in the database. They could also be used to create a competing expert system based on a different technique (e.g. neural network, discriminate analysis, etc) and comparing it to the existing system.

Clinical case data collection is a difficult task and requires careful planning which includes establishing criteria for study subject participation, completeness of records, and procedures for follow-up to confirm the initial diagnosis. These criteria will reduce case selection bias. Data would probably be more reliable if it was collected by medical practitioners because of the difficulty in performing and interpreting certain clinical tests (e.g. the elicitation of rebound tenderness takes considerable experience). Case data should be reduced to machine readable format daily and reviewed on a frequent basis. The inpatient and outpatient diagnoses should be coded by a medical records professional to reduce ambiguity. This would require an aggressive effort to maintain contact with clinicians, patients, and medical record personnel.

The data set presented suffered because of the large number of cases lost to follow up. Transcription errors may cast further doubt on the reliability of data, especially since the data set is small. Certainly one confirmed case of perforated duodenal ulcer is inadequate for testing. The number of cases could be increased by collaborating with other large medical centers, both civilian and government. Further involvement with multi-center groups interested in the development and testing of medical expert systems could provide additional case data.

References

1. Southerland DG. ABDX A Decision Support System for the Management of Acute Abdominal Pain Version 3.0 Programmer's Manual, NSMRL Report 1148, 1989.

2. Caras BC, Southerland DG, Fisherkelier KD. Medic-Abdominal Pain User's Manual, NSMRL Report 1146, 1989.

3. de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-Aided diagnosis of acut`e abdominal pain. British Medical Journal 1972;2.

4. de Dombal FT. Modification of the de Dombal computer-based system for diagnosis of abdominal pain to optimize its accuracy when applied to the United States Naval Submarine Population, Final report on U.S. Navy Contract No. N68171-79-C-9602.

5. Osborne S.F. Computer-assisted diagnosis of abdominal pain: Interim Report - July 1982-September 1983. NSMRL Report No. 1012, 1983.

APPENDIX A

Prior Probabilities of Disease used by the NSMRL Program

Dyspepsia is added to the non-specific abdominal pain category for reporting of results.

Disease	Prior Probability
Appendicitis (APPY)	0.18
Non-Specific Abdominal Pain (NSAP)	0.54
Renal Colic (RENC)	0.03
Perforated Duodenal Ulcer (PDU)	0.01
Cholecystitis (CHOL)	0.05
Small Bowel Obstruction (SBO)	0.03
Dyspepsia (DYSP)	0.16

Conditional Probabilities for Sign/Symptom Complex

Entries range from 0.1 to 99. The program will not accept conditional probabilities of 0 accounting for 0.1 values. The complete matrix is included although certain information (e.g., females, age >60) are not relevant to our study.

Sign or Symptom	Conditional Probability						
	APPY	NSAP	RENC	PDU	CHOL	SBO	DYSP
1. MALE	.1	.1	.1	.1	.1	.1	.1
2. FEMALE	.1	.1	.1	.1	.1	.1	.1
3. AGE 0-9	.1	.1	.1	.1	.1	.1	.1
4. AGE 10-19	25	19	05	08	.1	08	12
5. AGE 20-29	48	51	19	16	08	16	38
6. AGE 30-39	15	09	32	14	23	16	21
7. AGE 40-49	07	17	33	32	35	20	23
8. AGE 50-59	06	04	11	30	34	40	06
9. AGE 60-69	.1	.1	.1	.1	.1	.1	.1
10. AGE >69	.1	.1	.1	.1	.1	.1	.1
11. PAIN ONSET RUQ	03	01	.1	06	38	02	12
12. PAIN ONSET LUQ	01	03	.1	.1	02	02	08
13. PAIN ONSET RLQ	19	14	14	03	.1	02	.1
14. PAIN ONSET LLQ	02	09	11	.1	.1	02	.1

15. PAIN ONSET UPPER 1/2	10	20	01	59	45	28	58
16. PAIN ONSET LOW HALF	05	12	07	04	02	20	06
17. PAIN ONSET RT HALF	02	06	18	03	03	.1	03
18. PAIN ONSET LEFT HALF	01	04	08	.1	.1	.1	.1
19. PAIN ONSET CENTRAL	49	29	01	12	11	46	12
20. PAIN ONSET GENERAL	10	04	.1	14	.1	06	02
21. PAIN ONSET RT FLANK	.1	.1	18	.1	.1	.1	.1
22. PAIN ONSET LT FLANK	.1	01	26	.1	.1	.1	.1
23. NO PAIN AT ONSET	.1	.1	01	.1	.1	.1	.1
24. PAIN NOW RUQ	01	03	.1	02	42	.1	12
25. PAIN NOW LUQ	.1	02	.1	01	.1	.1	06
26. PAIN NOW RLQ	68	25	14	02	.1	02	.1
27. PAIN NOW LLQ	01	05	15	.1	.1	08	.1
28. PAIN NOW UPPER HALF	02	17	03	46	42	22	56
29. PAIN NOW LOWER HALF	07	12	05	01	02	14	02
30. PAIN NOW RIGHT HALF	04	03	14	11	02	.1	02
31. PAIN NOW LEFT HALF	.1	02	09	.1	.1	.1	.1
32. PAIN NOW CENTRAL	14	21	01	07	09	40	14
33. PAIN NOW GENERAL	03	03	.1	34	.1	14	03
34. PAIN NOW RT FLANK	01	01	20	.1	.1	.1	.1
35. PAIN NOW LT FLANK	01	01	26	.1	.1	.1	.1
36. NO PAIN NOW	.1	07	05	.1	03	.1	06
37. PAIN INTERMITTENT	05	18	14	05	02	.1	14
38. PAIN STEADY	80	46	34	86	73	22	61
39. PAIN COLICKY	15	36	52	09	25	80	25
40. PAIN IS MODERATE	63	50	11	05	27	38	44
41. PAIN IS SEVERE	37	50	89	95	73	62	56
42. MOVEMENT AGGRAVATES	53	24	17	48	09	18	18
43. COUGHING AGGRAVATES	22	09	.1	12	06	06	08
44. BREATHING AGGRAVATES	02	06	03	11	05	.1	05
45. FOOD AGGRAVATES	.1	03	.1	.1	11	02	06
46. AGGRAVATED BY OTHER	08	10	10	04	17	14	14
47. NOTHING AGGRAVATES	22	47	70	37	52	64	55
48. PROGRESS - BETTER	18	39	35	10	18	16	43
49. PROGRESS - SAME	30	35	26	44	49	50	39
50. PROGRESS - WORSE	52	25	39	46	33	34	18
51. DURATION <12 HRS	40	68	95	87	71	48	86
52. DURATION 12-24 H	60	32	05	13	29	52	14
53. DURATION 24-48 H	.1	.1	.1	.1	.1	.1	.1
54. DURATION 48+HRS	.1	.1	.1	.1	.1	.1	.1
55. LYING STILL RELIEVES	21	13	06	23	03	02	11
56. VOMITING RELIEVES	04	06	03	02	02	10	03
57. ANTACIDS RELIEVE	01	02	.1	03	.1	.1	05
58. FOOD RELIEVES	.1	01	.1	01	.1	.1	03
59. RELIEVED BY OTHER	12	12	26	06	21	20	26
60. NOTHING RELIEVES	63	66	65	65	74	68	58

61. NAUSEA PRESENT	65	62	72	61	68	81	65
62. NO NAUSEA	35	38	28	39	32	19	35
63. VOMITING PRESENT	55	42	70	54	72	86	62
64. NO VOMITING	45	58	30	46	28	14	38
65. BOWELS NORMAL	80	86	81	87	75	66	76
66. CONSTIPATION PRESENT	11	04	11	11	16	28	09
67. DIARRHEA PRESENT	09	08	08	02	08	06	11
68. BLOOD IN STOOLS	.1	01	.1	.1	01	.1	06
69. MUCUS IN STOOLS	.1	.1	.1	.1	.1	.1	.1
70. APPETITE DECREASED	70	37	40	48	62	66	46
71. APPETITE NORMAL	30	63	60	52	37	34	54
72. JAUNDICE PRESENT	.1	.1	02	.1	.1	.1	02
73. NO JAUNDICE	99	99	98	99	99	99	98
74. URINATION NORMAL	90	88	59	96	90	94	96
75. URINATION - FREQUENT	06	06	23	03	05	.1	02
76. URINATION - PAINFUL	04	06	18	01	03	04	.1
77. URINATION - DARK	.1	.1	04	.1	02	02	02
78. BLOOD IN URINE	.1	.1	01	.1	.1	.1	.1
79. PREVIOUS INDIGESTION	19	20	18	71	48	21	68
80. NO PREV. INDIGESTION	81	80	82	29	52	79	32
81. PREV. SIMILAR PAIN	37	40	46	47	75	57	67
82. NO PREV. SIM. PAIN	63	60	54	53	25	43	33
83. PREV. ABD. SURGERY	02	13	18	21	31	86	22
84. NO PREV. ABD. SURG.	98	87	82	79	69	14	78
85. PREVIOUS ILLNESS(es)	.1	.1	.1	.1	.1	.1	.1
86. NO PREVIOUS ILLNESS	.1	.1	.1	.1	.1	.1	.1
87. TAKING MEDS	10	11	18	25	35	31	44
88. NOT TAKING MEDS	90	89	82	75	65	69	56
89. TEMP <98.6	.1	.1	.1	.1	.1	.1	.1
90. TEMP 98.6 - 100.2	.1	.1	.1	.1	.1	.1	.1
91. TEMP 100.3 - 102	.1	.1	.1	.1	.1	.1	.1
92. TEMP >102	.1	.1	.1	.1	.1	.1	.1
93. PULSE <80	.1	.1	.1	.1	.1	.1	.1
94. PULSE 80-99	.1	.1	.1	.1	.1	.1	.1
95. PULSE >99	.1	.1	.1	.1	.1	.1	.1
96. SYST. BP <90	.1	.1	.1	.1	.1	.1	.1
97. SYST. BP 90-129	.1	.1	.1	.1	.1	.1	.1
98. SYST. BP >129	.1	.1	.1	.1	.1	.1	.1
99. DIAST. BP <70	.1	.1	.1	.1	.1	.1	.1
100. DIAST. BP 70-89	.1	.1	.1	.1	.1	.1	.1
101. DIAST. BP >89	.1	.1	.1	.1	.1	.1	.1
102. MOOD NORMAL	71	71	65	20	73	38	62
103. MOOD DISTRESSED	17	14	28	59	14	45	12
104. MOOD ANXIOUS	12	15	07	20	14	17	25
105. COLOR NORMAL	58	81	75	43	69	60	76
106. COLOR PALE	14	12	23	48	29	32	16

107. COLOR FLUSHED	28	07	02	05	02	05	06
108. COLOR JAUNDICED	.1	.1	.1	.1	.1	.1	02
109. COLOR CYANOTIC	.1	.1	.1	04	.1	03	.1
110. WBC < 8000	07	40	01	01	01	01	40
111. WBC 8 100-10 000	07	23	01	01	01	01	23
112. WBC 10 100-12 000	18	17	01	01	01	01	17
113. WBC 12 100-15 000	32	11	01	01	01	01	11
114. WBC >15 000	35	08	01	01	01	01	08
115. ABD INSPECT. NORMAL	87	99	96	39	89	82	92
116. VISIBLE PERISTALISIS	.1	.1	.1	.1	.1	06	.1
117. DECREASED ABD MOVE.	13	01	04	61	11	12	08
118. ABD SCARS PRESENT	02	14	20	21	28	84	23
119. NO ABDOMINAL SCARS	98	86	80	79	72	16	77
120. GUARDING PRESENT	72	24	23	62	62	38	30
121. NO GUARDING	28	76	77	38	38	62	70
122. RIGIDITY PRESENT	.1	.1	.1	.1	.1	.1	.1
123. NO RIGIDITY	.1	.1	.1	.1	.1	.1	.1
124. BOWEL SOUNDS NORMAL	.1	.1	.1	.1	.1	.1	.1
125. BOWEL SOUNDS ABSENT	.1	.1	.1	.1	.1	.1	.1
126. HYPER. BOWEL SOUNDS	.1	.1	.1	.1	.1	.1	.1
127. ABDOMEN DISTENDED	03	01	03	03	10	63	.1
128. NO ABD DISTENTION	97	99	97	97	90	37	99
129. MASS(es) PRESENT	01	01	05	01	11	12	.1
130. NO ABD MASSES	99	99	95	99	89	88	99
131. TENDERNESS RUQ	01	04	01	03	69	.1	11
132. TENDERNESS LUQ	.1	05	01	.1	.1	.1	03
133. TENDERNESS RLQ	87	29	15	03	02	06	02
134. TENDERNESS LLQ	02	11	12	01	02	06	.1
135. TENDER UPPER HALF	02	11	.1	30	14	14	61
136. TENDER LOWER HALF	02	11	02	01	.1	10	03
137. TENDER RIGHT HALF	07	06	11	10	03	02	06
138. TENDER LEFT HALF	.1	05	11	.1	.1	02	02
139. CENTRAL TENDERNESS	.1	03	.1	01	.1	14	02
140. GENERAL TENDERNESS	03	01	.1	55	02	40	08
141. TENDERNESS RT FLANK	01	01	19	01	.1	.1	.1
142. TENDERNESS LT FLANK	.1	01	23	.1	.1	.1	.1
143. NO TENDERNESS	.1	19	19	.1	11	10	12
144. MURPHY'S POSITIVE	01	03	.1	03	68	.1	02
145. MURPHY'S NEGATIVE	99	97	99	97	32	99	98
146. REBOUND PRESENT	80	23	05	54	10	33	18
147. NO REBOUND	20	77	95	46	90	66	82
148. RECTAL - NORMAL	57	72	91	80	92	91	98
149. RECTAL MASS	.1	01	.1	.1	02	.1	.1
150. LT RECTAL TENDERNESS	03	04	05	.1	.1	.1	.1
151. RT RECTAL TENDERNESS	27	14	.1	03	02	03	02
152. GEN. RECTAL TENDERNESS	12	09	04	17	04	06	.1

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 074-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF THE REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NSMRL REPORT 1181	5. MONITORING ORGANIZATION REPORT NUMBER(S) NA
6a. NAME OF PERFORMING ORGANIZATION Naval Submarine Medical Research Laboratory 6b. OFFICE SYMBOL *(If Applicable)*	7a. NAME OF MONITORING ORGANIZATION Naval Medical Research and Development Command
6c. ADDRESS *(City, State, Zip Code)* Box 900, Naval Submarine Base, NLON, Groton, CT 06349-5900	7b. ADDRESS *(City, State, Zip Code)* 8901 Wisconsin Ave., Bethesda, MD 20889-5606
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Same as 7a 8b. OFFICE SYMBOL *(If Applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, Zip Code)* Same as 7b	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 63706N; PROJECT NO. M0095; TASK NO 005; WORK UNIT ACCESSION NO. DN277023

11. TITLE *(Include Security Classification)*
Trial of a computer based program for the diagnosis of abdominal pain in males

12. PERSONAL AUTHOR(S)
P. L. Perrotta and D. M. Stetson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Interim	FROM _____ TO _____	92 September 02	20

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUB-GROUP	Computer based diagnosis; abdominal pain; males

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This report presents and evaluates data collected in 1988 in an effort to verify the NSMRL abdominal pain diagnosite program. Overall diagnostic accuracy of the program was found to be 69% comparede to the 80% accuracy rate of emergency room physicians. Sensitivity and specificity for distinguishing surgical from non-surgical cases was 56% and 85% respectively. Additional measures of performance of the abdominal pain program are presented along with the limitations of the data set and recommendations for future validation efforts.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
X UNCLASSIFIED/UNLIMITED SAME AS REPORT DTIC USERS	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Susan D. Monty, Publications Office	22b. TELEPHONE *(Include Area Code)* (203) 449-3967 22c. OFFICE SYMBOL